

David A. Kessler · Herbert Levine

Scaling solution in the large population limit of the general asymmetric stochastic Luria-Delbrück evolution process

Received: / Accepted:

Abstract One of the most popular models for quantitatively understanding the emergence of drug resistance both in bacterial colonies and in malignant tumors was introduced long ago by Luria and Delbrück. Here, individual resistant mutants emerge randomly during the birth events of an exponentially growing sensitive population. A most interesting limit of this process occurs when the population size N is large and mutation rates are low, but not necessarily small compared to $1/N$. Here we provide a scaling solution valid in this limit, making contact with the theory of Levy α -stable distributions, in particular one discussed long ago by Landau. One consequence of this association is that moments of the distribution are highly misleading as far as characterizing typical behavior. A key insight that enables our solution is that working in the fixed population size ensemble is not the same as working in a fixed time ensemble. Some of our results have been presented previously in shortened form [11]

Keywords Luria-Delbrück, mutants, growth, alpha-stable distribution

PACS 87.23.Kg, 05.40.-a

1 Introduction

In 1943, Luria and Delbrück [16] (LD hereafter) used a simple evolutionary model together with laboratory experiments to argue that bacterial resistance

David A. Kessler
Department of Physics, Bar-Ilan University, Ramat-Gan, IL52900 Israel
E-mail: kessler@dave.ph.biu.ac.il

Herbert Levine
Center for Theoretical Biological Physics, Rice University, Houston, TX 77096, USA
E-mail: Herbert.Levine@rice.edu

(to bacteriophage, in their specific case) arose via the selection of pre-existing mutations. Their model postulated a constant probability at which dividing normal cells would have a resistant daughter; these mutations were assumed to not affect the growth rate until the bacteria were challenged. The LD model and various extensions thereof have been used extensively in recent years to study the emergence of antibiotic-resistant microbes [14, 15] as well as chemotherapy-resistant cancer cells [4, 18, 13, 19, 6]. Similar models are used to study the generation of differentiated cells from stem-like precursors [3].

There has been considerable effort in the mathematics and statistical physics communities devoted to understanding the probability distribution that arises from the LD process [10, 5, 1, 20, 8, 9]. While early efforts focused on various quasi-deterministic versions of the model where both wild-type and mutant cells or alternatively just the wild-type are assumed to undergo continuous exponential growth, more recently interest has focussed on the fully stochastic model. A major achievement has very recently been achieved by Antal and Krapivsky [2] who solve for the generating function of the full joint distribution of the number of wild-types and mutants as a function of time, starting from a single wild-type individual. Also noteworthy is the solution by Angerer [1] for the probability distribution of the number of mutants given a fixed total population, N , in the case of the symmetric birth-only version of the stochastic model, where both wild-type and mutant have identical birth rates and cell death is excluded. However, it has proven difficult to obtain much physical insight from these solutions, and to relate these solutions to those of the simpler quasi-deterministic models. In particular, the remarkable fact that the distribution in these simpler models becomes [17] a universal α -stable (Lévy) distribution in the limit of small mutation rate μ , and large N such that $\mu N \gg 1$ has had no known analogue in the fully stochastic models.

Here, we introduce a new strategy to solve this model, taking advantage of the fact that for most cases of interest the population size N is very large and the mutation rate μ is very small. For example, a tumor may have upwards of 10^8 cells when it is first detected and the mutation rate, even if elevated above normal human mutation rates by genomic instability, would still be 10^{-6} or smaller (depending on exactly what range of mutations are assumed to be able to confer resistance). Note that μN is in general large, and hence a perturbative treatment would not be accurate. We show that if one focuses attention on the fixed N ensemble, one can derive an alternative generating function equation which leads immediately to a simple integral expression for the probability distribution. One can recover the asymptotic properties of this distribution for the general case where the resistant population may have a different fitness than the wild-type.

The outline of this work is as follows. To illustrate our methodology and to lay out the general features of the LD process, in particular that it gives rise to a distribution $P(m)$ for the number m of mutants which is heavy-tailed, we first present the special case of zero death rate. We show the general features of the quasi-deterministic model, including the fact that at large μN one obtains a universal α -stable distribution is actually valid for the fully stochastic version as well. Next, we consider cell death in addition to division

and mutation, and show that even for this case we recover for large μN the limiting Landau distribution (for equal birth rates) and its generalization for the asymmetric case. Finally, we compare our results to those of Antal and Krapivsky, who formally solve the same problem with a fixed time ensemble. Some of our results have appeared previously in abbreviated form [11].

2 The symmetric pure birth model

To illustrate our basic ideas, we start with the relatively simple case of no cell death (i.e. a pure birth/mutation process) and with equal birth rates for both the wild-type and mutant subpopulations. If we start from a fixed number N_0 of initial wild-type cells and wish to study the ensemble at fixed total population N , the number of events is just $N - N_0$. Rather than computing the probability function at fixed time, we instead use the event number as the independent variable. Then, $P_N(0)$ is just the probability that in the $N - N_0$ births since the beginning, there were no mutations, thus

$$P_N(0) = (1 - \mu)^{N - N_0}. \quad (1)$$

The master equation for $N > N_0$, $m > 0$ reads

$$P_{N+1}(m) = \frac{1}{N} \left[P_N(m-1)\mu(N-m+1) + P_N(m)(1-\mu)(N-m) + P_N(m-1)(m-1) \right]. \quad (2)$$

To derive what we will refer to as the scaling solution for the probability distribution, we will keep the leading order in N term for each order in μ . Later, we will see exactly when this is valid. It is easy to check that this involves keeping the coefficient of N^k for terms proportional to μ^k . Let us denote this coefficient in $P_N(m)$ as A_m^k , so that

$$P_N(m) \approx \sum_{k=0}^{\infty} A_m^k y^k; \quad y \equiv \mu N. \quad (3)$$

. Substituting this term into the master equation, we can check that the leading order in N condition (the $(\mu N)^k$ term is automatically satisfied, but matching the next order in N term (proportional to $\mu^k N^{k-1}$) leads to the non-trivial condition for $m, k \geq 1$:

$$(k+m)A_m^k = A_{m-1}^{k-1} - A_m^{k-1} + (m-1)A_{m-1}^k, \quad (4)$$

with

$$A_0^k = (-1)^k / k!; \quad A_m^0 = 0, \quad (m = 1 \dots \infty). \quad (5)$$

The former condition comes from our exact solution of $P_N(0)$, Eq. (1), above and the latter from the fact that $P_N(m)$ for all $m > 0$ vanishes if $\mu = 0$. To solve this equation, we define the generating function

$$F(x, y) \equiv \sum_{k=1, m=1} A_m^k x^m y^k, \quad (6)$$

in terms of which the generating function, $P_N(x)$, of the probability distribution is given by

$$P_N(x) = F(x, \mu N) + e^{-y}. \quad (7)$$

From the recursion relation Eq. (4) we easily obtain the differential equation

$$\left[y \frac{\partial}{\partial y} + x(1-x) \frac{\partial}{\partial x} \right] F(x, y) = -y(1-x)F + xy e^{-y}. \quad (8)$$

As already discussed in [11], we can solve this equation by first converting it to a homogeneous equation via the transformation $F = G - e^{-y}$, so that

$$\left[y \frac{\partial}{\partial y} + x(1-x) \frac{\partial}{\partial x} \right] G(x, y) = -y(1-x)G, \quad (9)$$

with $G(y=0) = 1$. The method of characteristics tells us to write $G(x, y) = G(x(y), y)$, choosing $dx/dy = x(1-x)/y$, so that the characteristic is

$$x = \frac{y}{\alpha + y}, \quad (10)$$

where α labels the characteristic. Then the equation reads

$$y \frac{dG}{dy} = -\frac{\alpha y}{\alpha + y} G, \quad (11)$$

with the solution

$$G = \left(1 + \frac{y}{\alpha} \right)^{-\alpha}; \quad \alpha = \frac{y(1-x)}{x}, \quad (12)$$

so that

$$F = (1-x)^{y(1-x)/x} - e^{-y}, \quad (13)$$

and thus

$$P_N(x) = (1-x)^{\mu N(1-x)/x}. \quad (14)$$

This result was derived initially by Lea and Coulson [14], for a model which neglected the stochastic nature of the birth-death dynamics for the wild-type population. In our derivation, going to the scaling limit of N large, $m \ll N$, has resulted in a distribution which has no N_0 dependence. Since when N_0 is large, the wild-type stochasticity is negligible, the fully stochastic treatment has to reduce to the Lea-Coulson result in the scaling limit.

One can numerically evaluate the Lea-Coulson distribution from the generating function via a contour integration to pick out $P_N(m)$:

$$P_N(m) = \frac{1}{2\pi i} \oint \frac{dz}{z^{m+1}} (1-z)^{\mu N(1-z)/z}. \quad (15)$$

This is more conveniently computed by moving the contour to hug the branch cut along the real axis $x > 1$:

$$P_N(m) \approx \frac{1}{\pi} \int_1^\infty dx x^{-m-1} (x-1)^{-\mu N(x-1)/x} \sin \frac{\pi \mu N(x-1)}{x}. \quad (16)$$

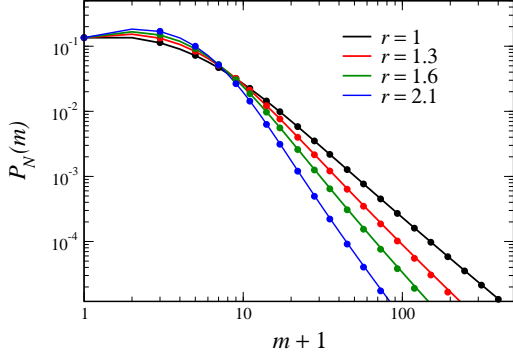


Fig. 1 The small μ , large N approximate distribution for $\mu N = 2$ for the pure birth-model, with the ratio of wild-type to mutant birth rate $r = 1, 1.3, 1.6$ and 2.1 (solid lines). This is shown together with the results (filled circles) of a direct numerical solution of the master equation for the parameters $\mu = 0.004$, $N = 500$.

A computation of this for the case $\mu N = 2$ is shown in Fig. 1, (labelled $r = 1$), together with results of a direct numerical solution of the master equation for the parameters $\mu = 0.004$, $N = 500$. The agreement is excellent even for this not so tiny value of μ .

It is clear by looking at the small x behavior that Eq. (14) correctly reproduces $P_N(0) \approx e^{-\mu N}$. In the small μN limit,

$$P_N(x) \approx 1 + \mu N \frac{1-x}{x} \ln(1-x) = 1 - \mu N \frac{1-x}{x} \sum_{m=1}^{\infty} \frac{x^m}{m}, \quad (17)$$

so that for $m \geq 1$

$$P_N(m) \approx \mu N \left(\frac{1}{m} - \frac{1}{m+1} \right) = \frac{\mu N}{m(m+1)}, \quad (18)$$

which is monotonically decreasing and has a fat-tailed $1/m^2$ decay. This $1/m^2$ behavior at large m is in fact true for all μN , and is a direct result of the logarithmic singularity of $P_N(x)$ at $x = 1$.

The large μN limit is also tractable, for m 's of order μN . We write $m = \mu N \tilde{m}$, and write $x = 1 + t/(\mu N)$, giving

$$P_N(m) = \frac{1}{\mu N \pi} \int_0^{\infty} dt e^{-\tilde{m}t - t \ln t / y} \sin \pi t = \frac{1}{\mu N \pi} \int_0^{\infty} dt e^{-(\tilde{m} - \ln y)t - t \ln t} \sin \pi t. \quad (19)$$

This is just a Landau distribution, the one-sided α -stable (Lévy) distribution with index $\alpha = 1$, written in terms of the original variables as $P_{\text{Landau}}(m/\mu N - \ln \mu N)$ times the normalization factor $1/\mu N$. The Landau distribution rises very rapidly from zero, reaches it maximum and then

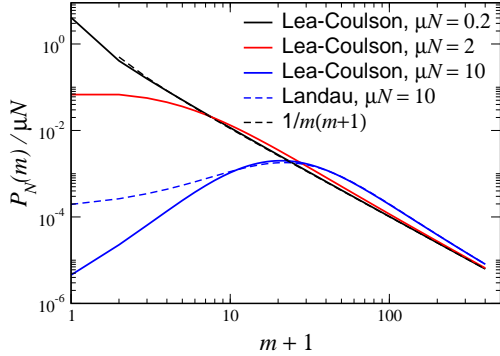


Fig. 2 The probability distributions $P_N(m)/\mu N$ as given by the Lea-Coulson formula for $\mu N = 0.2, 2$ and 10 . In addition, the small μN limiting form $1/m(m+1)$ and the large μN Landau distribution are also shown. It is seen that the Landau distribution correctly describes the distribution except for m 's of order unity.

falls with a $1/m^2$ fat tail. The appearance of the α -stable distribution in the quasi-deterministic model [17, 11] is a consequence of the fact that the total number of mutants is the sum of the mutants descended from each mutation event, and from large N these are independent and identically distributed with a power-law tail, and thus the conditions of the generalized central-limit theorem are satisfied. The Landau distribution arises in the fully stochastic model for the same reasons; the only difference being an overall shift in the distribution.

Thus, the Lea-Coulson distribution is seen to interpolate between the monotonic $\mu N/m(m+1)$ distribution at small μN to the fat-tailed one-sided α -stable distribution with its sharp rise for small m and $1/m^2$ falloff at large m . A signpost value is $\mu N = 2$, above which $P_N(m)$ is no longer monotonic. The position of the maximum moves rightward with increasing μN , asymptotically behaving as $m_{max} \approx \mu N (\ln \mu N - 0.22)$. This analysis is summarized in Fig. 2.

For m of order N , the Lea-Coulson result breaks down because we have neglected terms that become of comparable order in this new regime. This outer regime [11] describes how the distribution must drop to zero as we approach $m = N - N_0$; after all, the most extreme possibility is that every division resulted in adding one more mutant. Since μ is small, this part of the distribution arises when one of the first births lead to a mutant and hence the mutant sub-population has as much chance to expand exponentially as does the wild-type. To describe this scenario, one need only keep terms linear in μ . This can be done explicitly in the symmetric birth-only model [11]. However, even this problem is difficult to solve exactly in the general case, and is only of mathematical interest, since this outer part of the distribution describes extremely rare events and have essentially no physical relevance.

3 The non-symmetric pure birth problem

We now extend the previous results to account for a fitness (aka birth rate) difference between the wild-type and the resistant subpopulations. Normally it would be expected that alleles conferring resistance are less fit in the absence of any drug, and hence we expect the division rate ratio r of wild-type to resistant cell to be greater than one. The master equation connecting the mutant number distribution at one overall population size to the next overall population size, must take into account the relative probability that it is a wild-type cell that divides versus that it is a mutant cell which divides. If the current population has m mutant cells and hence $N - m$ wild-type ones, this relative probability is $(N - m)r / ((N - m)r + m)$. Then the equation clearly becomes

$$P_{N+1}(m) = (1 - \mu)P_N(m) \frac{(N - m)r}{(N - m)r + m} + \frac{P_N(m - 1)}{(N - m + 1)r + (m - 1)} [(N - m + 1)r\mu + (m - 1)]. \quad (20)$$

Again, we solve this equation in the limit of large N , small μ by keeping only the leading term in N for each power of μ , again using the expansion Eq. (3). Substituting this into the master equation, we obtain a generalization of Eq. (4):

$$(k + \frac{m}{r})A_m^k = A_{m-1}^{k-1} - A_m^{k-1} + \frac{(m - 1)}{r}A_{m-1}^k. \quad (21)$$

Following the same procedure as before, we transform this into an equation for the generating function,

$$\left[y \frac{\partial}{\partial y} + \frac{x(1 - x)}{r} \frac{\partial}{\partial x} \right] F(x, y) = -y(1 - x)F(x, y) + xy e^{-y}. \quad (22)$$

Again, the inhomogeneous solution is $-e^y$, and the homogeneous solution satisfies

$$\left[y \frac{\partial}{\partial y} + \frac{x(1 - x)}{r} \frac{\partial}{\partial x} \right] G(x, y) = -y(1 - x)G(x, y), \quad (23)$$

with $G(y = 0) = 1$. Again, we solve by the method of characteristics, with the characteristic labelled by α given by

$$x_\alpha(y) = \frac{y^{1/r}}{\alpha + y^{1/r}}, \quad (24)$$

so that

$$\frac{d}{dy} G(y, x(y)) = -\frac{\alpha}{\alpha + y^{1/r}} G, \quad (25)$$

with the solution

$$G = \exp \left(- \int_0^y \frac{\alpha}{\alpha + t^{1/r}} dt \right) = \exp \left(-y {}_2F_1 \left(1, r; 1 + r; -\frac{y^{1/r}}{\alpha} \right) \right), \quad (26)$$

where ${}_2F_1$ is a hypergeometric function. Substituting $\alpha = y^{1/r}(1-x)/x$, we get

$$G = \exp\left(-y {}_2F_1\left(1, r; 1+r; -\frac{x}{1-x}\right)\right) = \exp(-y(1-x) {}_2F_1(1, 1; 1+r; x)), \quad (27)$$

where we have employed a linear transformation identity, so that finally

$$P_N(x) = \exp(-\mu N(1-x) {}_2F_1(1, 1; 1+r; x)). \quad (28)$$

Again, to evaluate P_m we integrate around the branch cut, which again extends rightward from $x = 1$. Using standard identities involving hypergeometric functions, we have

$$\begin{aligned} {}_2F_1\left(1, r; 1+r; -\frac{x}{1-x}\right) &= \frac{r}{r-1} \left(\frac{1-x}{x}\right) {}_2F_1\left(1, 1-r; 2-r; -\frac{1-x}{x}\right) \\ &\quad + \frac{\pi r}{\sin \pi r} \left(\frac{1-x}{x}\right)^r {}_2F_1\left(r, 0; r; -\frac{1-x}{x}\right) \\ &= \frac{r}{r-1} \left(\frac{1-x}{x}\right) {}_2F_1\left(1, 1-r; 2-r; -\frac{1-x}{x}\right) \\ &\quad + \frac{\pi r}{\sin \pi r} \left(\frac{1-x}{x}\right)^r. \end{aligned} \quad (29)$$

This allows us to evaluate the discontinuity along the branch cut, giving

$$\begin{aligned} P_N(m) &= \frac{1}{\pi} \int_1^\infty \frac{dx}{x^{m+1}} \exp\left(\mu N \frac{r(x-1)}{(r-1)x} {}_2F_1\left(1, 1-r, 2-r, \frac{x-1}{x}\right)\right) \\ &\quad \times \exp\left(-\mu N \frac{\pi r}{\tan \pi r} \left(\frac{x-1}{x}\right)^r\right) \sin\left(\mu N \pi r \left(\frac{x-1}{x}\right)^r\right). \end{aligned} \quad (30)$$

Again, Fig. 1 presents a comparison of this expression with results of a direct numerical solution of the master equation for the parameters $\mu = 0.004$, $N = 500$. The agreement is again excellent.

The non-analyticity at $x = 1$ as before governs the large m asymptotics (subject to $m \ll N$, as usual). To investigate the behavior near $x = 1$, we use Eq. (29) to obtain ($t \equiv 1-x$)

$${}_2F_1\left(1, r; 1+r; -\frac{x}{1-x}\right) \approx \frac{r}{r-1} t + \frac{r\pi}{\sin \pi r} t^r. \quad (31)$$

By standard Tauberian theorems, this corresponds to a m^{1+r} decay at large m , independent at μN . In particular, for small μN , we have

$$\begin{aligned} P_N(x) &\approx 1 - \mu N(1-x) {}_2F_1(1, 1; 1+r; x) \\ &= 1 - \mu N + \mu N \sum_{m=1}^{\infty} \frac{(m-1)! \Gamma(1+r)r}{\Gamma(r+m+1)}, \end{aligned} \quad (32)$$

so that for $m \geq 1$:

$$P_N(m) \approx \mu N \frac{(m-1)! \Gamma(1+r)r}{\Gamma(r+m+1)} \xrightarrow{m \rightarrow \infty} \mu N \frac{r \Gamma(1+r)}{m^{1+r}}. \quad (33)$$

For large μN , as we saw in the case of $r = 1$, the contour integral giving $P_N(m)$ is dominated by the $x \approx 1$ singularity of the integrand. Thus,

$$P_N(m) \approx \int_0^\infty dt e^{-mt} \exp\left(-\mu N \left[\frac{r}{r-1}t + \frac{r\pi}{\sin \pi r} t^r\right]\right). \quad (34)$$

This is nothing but the one-sided α -stable distribution, now with index $\alpha = r$, consistent with the power-law decay. To see this, note that the characteristic function, ϕ (the Fourier transform of the probability density) of the α -stable distribution with index r has the form

$$\phi = \exp[it\nu - |ct|^r(1 - i \operatorname{sgn}(t) \tan(\pi r/2))], \quad (35)$$

where ν is called the shift parameter, which is just the mean of the distribution, and c is the scale parameter. By rotating $t \rightarrow -it$ in Eq. (34), we get for the shift parameter

$$\nu = \mu N \frac{r}{r-1}, \quad (36)$$

which is indeed the mean of the distribution Eq. (32). For the scale parameter, we get

$$c = \left[\mu N \frac{r\pi}{\sin \pi r} \cos \frac{\pi r}{2}\right]^{1/r} = \left[\mu N \frac{r\pi}{2 \sin \pi r/2}\right]^{1/r}. \quad (37)$$

4 General Asymmetric process with death

We now study the general asymmetric LD model with death, so that each event is either a birth event with rates b_w , b_m or a death event with rates d_w , d_m . We start from the probability distribution as a function of time. The master equations for $P(N, m, t)$, where N denotes the total (w.t. + mutated) population and m is the number of mutants, are:

$$\begin{aligned} \dot{P}(0, 0) &= d_w P(1, 0) + d_b P(1, 1), \\ \dot{P}(N, m) &= b_w(N-1-m)(1-\mu)P(N-1, m) + b_w \mu(N-m)P(N-1, m-1) \\ &\quad + b_m(m-1)P(N-1, m-1) \\ &\quad + d_w(N+1-m)P(N+1, m) + d_m(m+1)P(N+1, m+1) \\ &\quad - [(b_w + d_w)(N-m) + (b_m + d_m)m]P(N, m); \end{aligned} \quad (N > 1), \quad (38)$$

with $P(N_0, 0) = 1$ at $t = 0$.

To simplify these expressions, we make use of the fact that we are interested in a fixed value of N , independent of how much time has elapsed since the beginning. From the biology perspective, what one knows (ideally) is the

size of the tumor at discovery, with no idea how long the tumor has been growing. Thus, we are interested in the conditional probabilities

$$P(m|N) \equiv (Z_N)^{-1} \int_0^\infty P(N, m; t) dt \equiv (Z_N)^{-1} T_N(m); \quad Z_N \equiv \sum_m T_N(m). \quad (39)$$

Integrating the master equation over time leads to the following equation for $T_N(m)$, $N \geq 1$:

$$\begin{aligned} \left[(b_w + d_w)(N - m) + (b_m + d_m)m \right] T_N(m) - \delta_{N, N_0} \delta_{m, 0} \\ = b_w(N - 1 - m)(1 - \mu)T_{N-1}(m) \\ + b_w\mu(N - m)T_{N-1}(m - 1) + b_m(m - 1)T_{N-1}(m - 1) \\ + d_w(N + 1 - m)T_{N+1}(m) + d_m(m + 1)T_{N+1}(m + 1). \end{aligned} \quad (40)$$

with the definition $T_0(m) \equiv 0$. Again we assume $T_N(m) = \sum_k (\mu N)^k A_m^k$. Then, plugging this into the recursion relation, and noting that as always the N^k term cancels, we find the generalized scaling-limit recursion

$$\begin{aligned} A_m^k [m(b_m + d_m) + k(b_w - d_w)] = b_w A_{m-1}^{k-1} - b_w A_m^{k-1} + b_m(m - 1)A_{m-1}^k \\ + d_m(m + 1)A_{m+1}^k. \end{aligned} \quad (41)$$

By now, the procedure for solving this equation should be obvious. We rewrite this systems as an equation for the generating function'

$$\left[(b_w - d_w)y \frac{\partial}{\partial y} + (x(b_m + d_m) - b_m x^2 - d_m) \frac{\partial}{\partial x} \right] F(x, y) = -b_w y(1 - x)F + \mathcal{I}, \quad (42)$$

where \mathcal{I} is the inhomogeneous term. Again $x = 1$ is the singular point with the coefficient of $\partial F / \partial x$ vanishes. Defining $r \equiv (b_w - d_w)/(b_m - d_m)$, the ratio of *net* birth rates, the characteristic is given by

$$x(y) = \frac{y^{1/r} + \alpha d_m}{y^{1/r} + \alpha b_m}. \quad (43)$$

The homogeneous equation now reads

$$(b_w - d_w)y \frac{dG}{dy} = -b_w y \frac{\alpha(b_m - d_m)}{y^{1/r} + \alpha b_m} G. \quad (44)$$

Notice that except for factors, this is the exact same equation we encountered in the asymmetric pure-birth case, and so the solution is immediate:

$$\begin{aligned} G &= \mathcal{N} \exp \left(-\frac{b_w}{r b_m} y {}_2F_1 \left(1, r; 1 + r; -\frac{y^{1/r}}{\alpha} \right) \right) \\ &= \mathcal{N} \exp \left(-y \frac{b_w}{r b_m} {}_2F_1 \left(1, r; 1 + r; -\frac{b_m x - d_m}{b_m(1 - x)} \right) \right), \end{aligned} \quad (45)$$

where \mathcal{N} is a normalization prefactor. Our desired conditional probability generating function, which we denote by $P_N(x)$ is then

$$P_N(x) = \exp \left(-\mu N \frac{b_w}{rb_m} {}_2F_1 \left(1, r; 1+r; -\frac{b_mx - d_m}{b_m(1-x)} \right) \right). \quad (46)$$

We see that in the pure birth case, $d_w = d_m = 0$, this reproduces the results of the previous section.

The large m behavior is, as we have seen, determined by the behavior in the vicinity of $x = 1$. Writing $x = 1 - t$, we have

$${}_2F_1 \left(1, r; 1+r; -\frac{b_mx - d_m}{b_m(1-x)} \right) \approx \frac{rb_m}{(r-1)(b_m - d_m)} t + \frac{r\pi}{\sin \pi r} \left(\frac{b_m}{b_m - d_m} t \right)^r. \quad (47)$$

The analytic structure is identical to the asymmetric pure-birth case, and so $P(m|N)$ again decays as m^{1+r} :

$$P(m|N) \approx \mu N \left(\frac{b_m}{b_m - d_m} \right)^r \frac{b_w \Gamma(1+r)}{b_m m^{1+r}}. \quad (48)$$

Though the exponent is only a function of the the ratio of the *net* birth rates, the coefficient is more complicated.

We can again evaluate $P(m|N)$ from our generating function via an integration along the branch cut:

$$P_N(m) = \frac{1}{\pi} \int_1^\infty \frac{dx}{x^{m+1}} \exp \left(\mu N \frac{b_w(x-1)}{(r-1)(b_mx - d)} {}_2F_1 \left(1, 1-r, 2-r, \frac{b_m(x-1)}{b_mx - d} \right) \right) \\ \times \exp \left(-\mu N \frac{\pi b_w}{b_m \tan \pi r} \left(\frac{b_m(x-1)}{b_mx - d} \right)^r \right) \sin \left(\mu N \frac{\pi b_w}{b_m} \left(\frac{b_m(x-1)}{b_mx - d} \right)^r \right). \quad (49)$$

We present in Fig. 3 such a calculation for the case $b_w = 3/4$, $d_w = 1/4$, $b_m = b_w/r$, $d_m = d_w/r$, with $r = 1.3$. Also shown is a direct numerical solution of the master equation, Eq. (40), for $\mu = 0.004$, $N = 500$, which confirms our asymptotic analysis. For comparison, the pure birth distribution with the same $r = 1.3$ is shown.

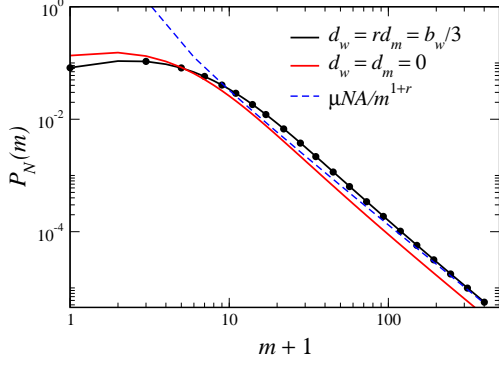


Fig. 3 The generalized Lea-Coulson distribution $P(m|N)$ for the asymmetric birth-death model, together with a direct solution of the master equation, Eq. (40) for $\mu = 0.004$, $N = 500$. Here $b_w = 3/4$, $d_w = 1/4$, $b_m = b_w/r$, $d_m = d_w/r$, with $r = 1.3$. The asymptotic power law, Eq. (48) is also shown. For comparison, the pure birth distribution with the same $r = 1.3$ is shown.

For small μN , we can write $P(m|N)$ explicitly, since

$$\begin{aligned}
 P_N(x) &\approx 1 - \mu N \frac{b_w}{rb_m} {}_2F_1\left(1, r; 1+r; -\frac{b_mx - d_m}{b_m(1-x)}\right) \\
 &= 1 - \mu N \frac{b_w}{rb_m} \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{d_m}{b_m(1-x)}\right)^k \frac{\Gamma(1+k)\Gamma(r+k)\Gamma(1+r)}{\Gamma(1)\Gamma(r)\Gamma(1+r+k)} \\
 &\quad \times {}_2F_1\left(1+k, r+k; 1+r+k; -\frac{x}{1-x}\right) \\
 &= 1 - \mu N \frac{b_w}{rb_m} \sum_{k=0}^{\infty} \left(\frac{d_m}{b_m}\right)^k \frac{r}{r+k} (1-x) {}_2F_1(1, 1+k; 1+r+k; x) \\
 &= 1 - \mu N \frac{b_w}{b_m} \sum_{k=0}^{\infty} \left(\frac{d_m}{b_m}\right)^k \frac{1-x}{r+k} \sum_{m=0}^{\infty} \frac{\Gamma(1+m)\Gamma(1+k+m)\Gamma(1+r+k)}{\Gamma(1)\Gamma(1+k)\Gamma(1+r+k+m)m!} x^m \\
 &= 1 + \mu N \frac{b_w}{b_m} \sum_{k=0}^{\infty} \left(\frac{d_m}{b_m}\right)^k \left[-\frac{1}{r+k} + \sum_{m=1}^{\infty} \frac{r\Gamma(k+m)\Gamma(r+k)}{\Gamma(1+k)\Gamma(1+r+k+m)} x^m \right] \\
 &= 1 + \mu N \frac{b_w}{rb_m} \left[-{}_2F_1\left(1, r; 1+r; \frac{d_m}{b_m}\right) \right. \\
 &\quad \left. + \sum_{m=1}^{\infty} \frac{r^2\Gamma(r)\Gamma(m)}{\Gamma(r+m+1)} {}_2F_1\left(r, m; r+m+1; \frac{d_m}{b_m}\right) x^m \right],
 \end{aligned} \tag{50}$$

so that, for $m > 0$,

$$P(m|N) \approx \mu N r \frac{b_w}{b_m} \frac{\Gamma(r)\Gamma(m)}{\Gamma(r+m+1)} {}_2F_1\left(r, m; r+m+1; \frac{d_m}{b_m}\right) \equiv p(m|N), \quad (51)$$

and

$$P(0|N) \approx 1 - \mu N \frac{b_w}{r b_m} {}_2F_1\left(1, r; 1+r; \frac{d_m}{b_m}\right) \equiv 1 + p(0|N). \quad (52)$$

It is straightforward, though tedious, to show that

$$t_N(m) = \frac{1}{b_w N} p(m|N) \quad (53)$$

is an *exact* solution of the small μ version of Eq. (40) (using the fact that to zeroth order in μ , $T_N(0) \approx 1/(b_w N)$):

$$\begin{aligned} [(b_w + d_w)(N - m) + (b_m + d_m)m]t_N(m) &= b_w(N - 1 - m)t_{N-1}(m) + b_m(m - 1)t_{N-1}(m - 1) \\ &\quad + d_w(N + 1 - m)t_{N+1}(m) + d_m(m + 1)t_{N+1}(m + 1), \\ [(b_w + d_w)(N - 1) + (b_m + d_m)]t_N(1) &= b_w(N - 2)t_{N-1}(1) + d_w N t_{N+1}(1) + 2d_m t_{N+1}(2) + \mu, \\ [(b_w + d_w)N]t_N(0) = b_w(N - 1)t_{N-1}(0) + d_w(N + 1)t_{N+1}(0) &+ d_m t_{N+1}(1) - \mu, \end{aligned} \quad (54)$$

with $0 \leq m \leq \infty$. However, the finite N system is truncated, with the equation for $t_N(N)$ being different, so that this solution does not represent an exact solution of this truncated system. Nevertheless, it turns out that $t_N(0)$ and $t_N(N)$ decouple from the other equations, and in fact $T_N(m) = t_N(m)$ is the exact solution of the $N_0 = 1$ small μ problem for $0 < m < N$.

The perspicacious reader will have noticed that the problem we are solving in this section does not reduce exactly to the problem we posed when we considered the pure-birth asymmetric model. The equations are the same with the identification that $T_N(m)$ in our section maps to $P_N(m)/((N - m)r + m)$ in that section. However, the probabilities we construct from the T_N do not correspond to the P 's in the previous section. Here, we weight each state by the time spent in the state, which we did not do there. To leading order in N , however, the time spent in a state is independent of the state (as long as $m \ll N$), so the large- N solution is identical in the two ensembles.

Note that when $x = 0$, we get

$$P(0|N) = \exp\left(-\mu N \frac{b_w}{r b_m} {}_2F_1\left(1, r; 1+r; \frac{d_m}{b_m}\right)\right), \quad (55)$$

which is equivalent to a well-known result of Ref. [8]. The advantage of our approach is that it is clear what approximations have to be made to obtain this result, something that is not overly clear in this earlier reference. It should also be noted in this context that the result obtained in Ref. [8] for $P(m|N)$ in the small μN limit is not correct.

We can use Eq. (47) to find the large μN , $m \sim \mu N$ limiting distribution, where the branch cut integral is as usual dominated by the singular point at $y = 1$. Given that the analytic structure here, as we have seen, is exactly the same as for the asymmetric pure-birth case, we immediately can conclude that the distribution for large μN reproduces the one-sided α -stable distribution with index r . The shift parameter is

$$\begin{aligned}\nu &= \mu N \frac{b_w}{b_m} \frac{b_m - d_m}{b_w - d_w} \frac{b_w - d_w}{b_w - d_w - b_m + d_m} \frac{b_m}{b_m - d_m} \\ &= \mu N \frac{b_w}{b_w - d_w - b_m + d_m}.\end{aligned}\quad (56)$$

That this is correct can be seen from the calculation of \bar{m} in the original DL model:

$$\begin{aligned}\bar{m} &= \int_0^T \mu b_w N_0 e^{(b_w - d_w)t} e^{(b_m - d_m)(T-t)} dt \\ &= \mu(N - N_0) \frac{b_w}{b_w - d_m - b_m + d_m}.\end{aligned}\quad (57)$$

For the scale parameter, we get

$$\begin{aligned}c &= \frac{b_m}{b_m - d_m} \left[\mu N \left(\frac{b_w}{b_m} \right) \left(\frac{b_m - d_m}{b_w - d_w} \right) \frac{b\pi}{\sin \pi b} \cos \frac{\pi b}{2} \right]^{1/b} \\ &= \left(\frac{b_m}{b_m - d_m} \right)^{1-1/b} \left[\mu N \left(\frac{b_w}{b_w - d_w} \right) \frac{b\pi}{2 \sin \frac{\pi b}{2}} \right]^{1/b}.\end{aligned}\quad (58)$$

Lastly, for completeness, we write down the results for the symmetric birth-death process. Here $r = 1$ and the large μN limit is thus again a Landau distribution. For $P(0|N)$ we have

$$P(0|N) \equiv e^{-\mu N b_w / d_w \ln \frac{b_w}{b_w - d_w}}. \quad (59)$$

For the large μN limit, we have

$$P(m|N) \approx \frac{b-d}{\pi \mu N b} \int_0^\infty dt e^{-tm/(\mu N b/(b-d) - t \ln t + t \ln \mu N)} \sin \pi t, \quad (60)$$

which is again a Landau distribution, $P_{\text{Landau}}(m(b-d)/\mu N b - \ln \mu N)$. Amazingly, introducing death to the LD problem hardly makes a change in the answer, just renormalizing the variable m by the ratio $(b-d)/b$.

5 Comparison with the fixed time ensemble

5.1 Pure birth model

The most general solution of LD problem is given in the recent paper by Antal and Krapivsky (AK) [2], who solve the problem of determining the probability distribution at a fixed time. We first present their solution, modifying their notation slightly so as not to interfere destructively with the notation we have been using until now. They write down the backward version of the master equation for the probabilities of w wild-types and m mutants in the form

$$\begin{aligned} \frac{dP^w}{dt}(w, m, t) &= b_w(1 - \mu)P^{ww}(w, m, t) + \mu b_w P^{wm}(w, m, t) \\ &\quad - (b_w + d_w)P^w(w, m, t) + d_w \delta_{w,0} \delta_{m,0}, \\ \frac{dP^m}{dt}(w, m, t) &= b_m P^{mm}(w, m, t) - (b_m + d_m)P^m(w, m, t) + d_m \delta_{N,0} \delta_{m,0}. \end{aligned} \quad (61)$$

In this expression the division and death rates of the wild-type (w) and mutant (m) subpopulations are denoted by b and d with the proper subscript. The superscript denotes a process which starts from a single wild-type individual (superscript w), a single mutant individual (superscript m), two wild-type (ww) and one of each (wm). A great simplification occurs in passing to the generating function version, in that the independent time evolution of any two individuals leads to factorization of the two-individual probabilities, so that the generating function for P^{ww} is the square of the generating function for (P^w), with similar results for P^{wm} and P^{ww} . This then leads to a pair of coupled ODE's for $W(x, y)$ and $M(x, y)$, the generating functions defined as

$$W(x, y; t) \equiv \sum x^m y^w P^w(w, m; t); \quad M(x, y; t) \equiv \sum x^m y^w P^m(w, m; t), \quad (62)$$

and the equations are

$$\begin{aligned} \frac{\partial W}{\partial t} &= b_w(1 - \mu)W^2 + \mu b_w MW + d_w - (b_w + d_w)W, \\ \frac{\partial M}{\partial t} &= b_m M^2 + d_m - (b_m + d_m)M. \end{aligned} \quad (63)$$

To see how the solution of this system works and how it relates to our approach, we return for the moment to the simpler case of no death and symmetric division $b_w = b_m \equiv b$, and recall that the AK derivation so far is for $N_0 = 1$. We can rewrite these equations by picking the unit of time to set the wild-type growth rate $b(1 - \mu)$ equal to 1. This leads to the new system

$$\begin{aligned} \frac{\partial W}{\partial t} &= W^2 + \nu MW - (1 + \nu)W \\ \frac{\partial M}{\partial t} &= \lambda M^2 - \lambda M \end{aligned} \quad (64)$$

with $\nu \equiv \mu/(1-\mu)$, $\lambda \equiv 1/(1-\mu)$. The initial condition for the M equation is that $M(t=0) = x$, since by definition the state defining M starts with one mutant cell. The solution of this equation is

$$M = -\frac{z}{1-z}; \quad z = -\frac{x}{1-x}e^{-\lambda t}. \quad (65)$$

Substituting this into the W equation, one obtains the solution

$$W = -\frac{Cz}{(1-z)^\mu + C - Cz}. \quad (66)$$

The constant of integration C is picked to satisfy the initial condition $W(t=0) = y$, yielding

$$C = -\frac{y(1-x)^{1-\mu}}{y-x}. \quad (67)$$

From this, one can obtain directly the distribution of mutant numbers, independent of the number of wild-type, given by $F(x, t) \equiv W(x, y=1; t)$; setting $y=1$ gives $C=1$ and using the mean population size $\bar{N} \equiv e^{\lambda t}$, we obtain

$$F(x) = \frac{x}{x + \bar{N}(1-x) [1 - (1-x + x/\bar{N})^\mu]}. \quad (68)$$

To compare this result with our finding given above in Eq. 13, we go to the scaling limit, which is small μ , large \bar{N} , finite m (so that the number of wild-type is approximately equal to the total population). We can then drop the x/\bar{N} term, and we have

$$F(x) \approx \frac{x}{x - (1-x)\mu\bar{N}\ln(1-x)}, \quad (69)$$

so that indeed $\mu\bar{N}$ is the relevant control parameter. This is *not* the Lea-Coulson generating function, as has already been noted in our abbreviated initial paper [11]. In particular, $P_0 = 1/(1 + \mu\bar{N})$, so that it does not decay exponentially with $\mu\bar{N}$ as we found above. The difference lies in the fact that the AK calculation is performed for *fixed time*, not at *fixed total population* N as before. For large time, the distribution of the number of wild-type cells, independent of the number of mutants can be shown to be geometric:

$$P^w(w) \approx \frac{1}{\bar{w}(t)} e^{-w/\bar{w}(t)}; \quad \bar{w}(t) = e^{b(1-\mu)t}. \quad (70)$$

From this, given that the probability of zero mutants given w wild-type cells is, for small μ , $\exp(-\mu(w-1))$, it immediately follows that the probability of zero mutants is given by

$$P_0 \approx \int_0^\infty \frac{dw}{\bar{w}(t)} e^{-\mu w} e^{-w/\bar{w}(t)} \approx \frac{1}{1 + \mu\bar{N}}, \quad (71)$$

in agreement with the AK calculation. Clearly, then, the configurations at fixed time with small numbers of wild-types and thus a relatively large probability of never having spawned a resistant cell increase P_0 drastically. In Fig.

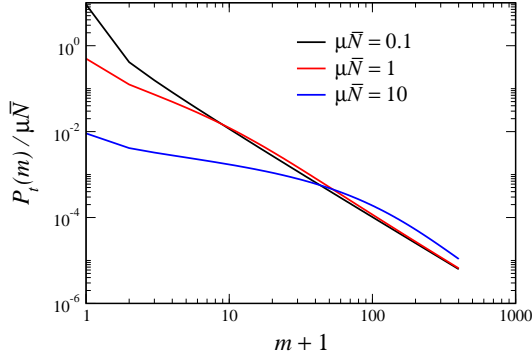


Fig. 4 The scaled distribution of mutants, $P_t(m)/\mu\bar{N}$, in the AK fixed time ensemble, for times corresponding to $\mu\bar{N} = 0.1, 1$ and 10 .

4 we present the distribution resulting from the AK fixed time ensemble, for times corresponding to $\mu\bar{N} = 0.1, 1$ and 10 . We see that in all cases the distribution is monotonic, and looks nothing like the Lea-Coulson result, nor does it go over to the Landau distribution at large $\mu\bar{N}$.

Looking at $F(x)$ for small $\mu\bar{N}$, we see that it does reproduce the small μN limit of the Lea-Coulson answer. More generally, the structure of the singularity at $x = 1$ is the same of for Lea-Coulson, so that fixed-time ensemble does show the same $1/m^2$ behavior as our fixed- N ensemble for the intermediate asymptotic regime $1 \ll m \ll N$. The behavior in the cut-off regime $m \sim N$ is not surprisingly very different from the fixed- N behavior, since the shift of the pole of $F(x)$ to $x = 1/(1 + 1/\bar{N})$ means that the decay of $P(m)$ is exponential for $m \sim N$.

Of course, the full AK solution contains information about the wild-type distribution encoded in its dependence on y which has been dropped in the derivation of F . If we return to the full solution for W , we can of course restrict ourselves to a fixed value of $m = \bar{N}$ by finding the coefficient of the term that goes like y^N . This is still not exactly the same problem we solved. Our initial ensemble allows the elapsed time to vary, whereas this new result fixes both the time and the total population. However, the distribution of mutants as we have seen is controlled by the total population, and we do not expect the time taken to reach this population to matter significantly, at least as long as it is close to the average time. Further, as we have already noted, for $m \ll N$, the total population N and the wild-type population w are the same to leading order in N . Thus, undaunted, we continue. The projection of $W(x, y)$ onto fixed $m = \bar{N}$, which we denote $W_{\bar{N}}(x)$ can be done exactly, since the y dependence of W is simple and takes the form

$$W = A_1 y / (A_2 y + A_3) \quad (72)$$

with

$$A_1 = z, \quad A_2 = z - 1 + \frac{[(1-z)(1-x)]^\mu}{1-x}, \quad A_3 = -\frac{x[(1-z)(1-x)]^\mu}{1-x}, \quad (73)$$

and where as before $z = -\frac{x/\bar{N}}{1-x}$. Thus,

$$W_{\bar{N}}(x) = \mathcal{N} \frac{A_1}{A_3} \left(-\frac{A_2}{A_3} \right)^{\bar{N}-1}. \quad (74)$$

where \mathcal{N} is a normalization factor, chosen so that $W_{\bar{N}}(1) = 1$. The non-trivial normalization factor arises because we do not know a priori the overall probability of finding exactly the correct value of N in the overall ensemble, and hence this must be divided out to obtain a meaningful probability distribution. Now, in the limit $x \rightarrow 1$,

$$z \rightarrow -\frac{1}{\bar{N}(1-y)} \rightarrow -\infty, \quad A_2 \rightarrow -\frac{1}{\bar{N}(1-y)} + \frac{\bar{N}^{-\mu}}{1-y} = \frac{\bar{N}^{-\mu}(1-\bar{N}^{\mu-1})}{1-y},$$

$$A_3 \rightarrow -\frac{\bar{N}^{-\mu}}{1-y}. \quad (75)$$

Thus, we can find the value of the normalization constant

$$\mathcal{N} = \bar{N}^{1-\mu} \left(1 - \bar{N}^{\mu-1} \right)^{1-\bar{N}} \approx \bar{N}^{1-\mu} e^{1+\mu \ln \bar{N}} = \bar{N} e^1. \quad (76)$$

Given this, we can now calculate the leading large- \bar{N} behavior of $W_{\bar{N}}(x)$, which we wish to compare to $F(x)$. For fixed x and large \bar{N} , $z \rightarrow 0$ and

$$A_2/A_3 = \frac{[(1-z)(1-x)]^{1-\mu}}{x} - \frac{1}{x} \approx -1 + \frac{1}{\bar{N}} - \mu \frac{(1-x) \ln(1-x)}{x}. \quad (77)$$

Also,

$$A_1/A_3 = -\frac{z(1-x)}{x[(1-z)(1-x)]^\mu} \approx \frac{1}{\bar{N}}. \quad (78)$$

. Finally we obtain

$$W_{\bar{N}}(x) \approx \bar{N} e^1 \left(\frac{1}{\bar{N}} \right) e^{-1+\mu \bar{N} \frac{(1-x) \ln(1-x)}{x}} = (1-x)^{\mu \bar{N} (1-x)/x}. \quad (79)$$

This is just the Lea-Coulson result that we had obtained above. Thus, our expectations that post-projecting the fixed-time ensemble for fixed wild-type abundance would reproduce the fixed population ensemble result in the intermediate asymptotic regime are indeed realized.

The above calculation was for $N_0 = 1$. For general N_0 , the AK result for the generating function $W(x, y; N_0)$ is immediate:

$$W(x, y; N_0) = [W(x, y)]^{N_0}, \quad (80)$$

due to the independence of the life histories of the initial N_0 wild-type. Clearly then, the AK result, i.e. conditioning on fixed time, is N_0 dependent, as opposed to the N_0 independence of the fixed total population result, for $m \ll N$. We can generalize our calculation of the fixed $m = \bar{N}$ projection of the AK result to general $N_0 \ll \bar{N}$. Now, we have $\exp(\lambda t) = \bar{N}/N_0$. Then, performing the Taylor expansion of W^{N_0} we get

$$W_{\bar{N}}(x; N_0) = \mathcal{N} \left(\frac{A_1}{A_3} \right)^{N_0} \binom{\bar{N}-1}{N_0-1} \left(-\frac{A_2}{A_3} \right)^{\bar{N}-N_0}. \quad (81)$$

The normalization factor is now

$$\begin{aligned} \mathcal{N} &= \left(\frac{\bar{N}}{N_0} \right)^{N_0(1-\mu)} \bar{N}^{1-N_0} (N_0-1)! \left[1 - \left(\frac{\bar{N}}{N_0} \right)^{\mu-1} \right]^{N_0-\bar{N}} \\ &\approx \bar{N} \left(\frac{e}{N_0} \right)^{N_0} (N_0-1)!. \end{aligned} \quad (82)$$

Then,

$$\begin{aligned} W_{\bar{N}}(x; N_0) &= \mathcal{N} \left(\frac{N_0}{\bar{N}} \right)^{N_0} \binom{\bar{N}-1}{N_0-1} \left(1 - \frac{N_0}{\bar{N}} + \mu \frac{(1-x) \ln(1-x)}{x} \right)^{\bar{N}-N_0} \\ &\approx (1-x)^{\mu \bar{N}(1-x)/x}, \end{aligned} \quad (83)$$

recovering, as we expect, the N_0 -independent Lea-Coulson result. It is remarkable that the fixed population ensemble results are much more similar to the original semi-deterministic model of Luria and Delbrück, with its universal aspects such as N_0 independence and large μN α -stable distribution than is the fixed time ensemble.

5.2 General Asymmetric Birth-Death Model

We chose to first present our projection method in the simplest possible context of the pure-birth, symmetric model, as in this case the AK solution is expressible in terms of elementary functions and the structure of the calculation is most transparent. The procedure, however, works in essentially the same way for the most general asymmetric birth-death model for which AK presented their solution for the fixed time ensemble. The model is now parameterized in terms of wild-type and mutant birth rates, b_w , b_m as well as wild-type and mutant death rates, d_w , d_m . AK choose to scale time so that the non-mutating wild-type birth process $W \rightarrow WW$ has rate unity. AK then introduce the following rates:

$$\begin{aligned} \alpha_m &\equiv \frac{b_m}{b_w(1-\mu)}, & \beta_w &\equiv \frac{d_w}{b_w(1-\mu)}, & \beta_m &\equiv \frac{d_m}{b_w(1-\mu)}, & \nu &= \frac{\mu}{1-\mu}, \\ \lambda_w &= 1 - \beta_w - \nu, & \lambda_m &= \alpha_m - \beta_m. \end{aligned} \quad (84)$$

AK also introduce the auxiliary parameter

$$\omega \equiv -\frac{\lambda_w}{2} - \frac{\nu\beta_m}{2\alpha_m} + \sqrt{\left(\frac{\lambda_w}{2} + \frac{\nu\beta_m}{2\alpha_m}\right)^2 + \frac{\nu\lambda_m}{\alpha_m}}. \quad (85)$$

Note that in the small μ limit, ω is of order μ :

$$\omega \approx \mu \frac{b_m - d_m}{b_w - d_w} \frac{b_w}{b_m}. \quad (86)$$

AK define as well

$$a \equiv \frac{\omega}{\lambda_m} \approx \frac{\mu b_w^2}{b_m(b_w - d_w)}, \quad b \equiv \frac{\omega + 1 - \beta_w}{\lambda_m} \approx \frac{b_w - d_w}{b_m - d_m},$$

$$c \equiv 1 + a + b - \frac{\nu}{\alpha_m} \approx b + 1, \quad (87)$$

where the approximations are correct to leading order in μ . Then, the full solution derived in AK takes the extremely complicated form

$$W(x, y; t) = 1 + \omega + \frac{T_1 y + T_2}{T_3 y + T_4} = 1 + \omega + \frac{(T_1 y + T_2)}{T_4} \sum_k (-1)^k \left(y \frac{T_3}{T_4} \right)^k, \quad (88)$$

where

$$\begin{aligned} T_1 &= z^{c+1} z_0 F_3(z) F_2(z_0) + (c-1) z_0^c F_2(z) F_1(z_0) z - z^2 z_0^c F_4(z) F_1(z_0), \\ T_2 &= z^{c+1} z_0 F_3(z) [F_2(z_0)(\lambda_m(c-1) - 1 - \omega) - z_0 F_4(z_0) \lambda_m] \\ &\quad - z_0^c z (c-1) F_2(z) [F_1(z_0)(1 + \omega) + \lambda_m F_3(z_0) z_0] \\ &\quad + z_0^c z^2 F_4(z) [F_1(z_0)(1 + \omega) + \lambda_m z_0 F_3(z_0)], \\ T_3 &= z^c z_0 F_1(z) F_2(z_0) - z_0^c z F_2(z) F_1(z_0), \\ T_4 &= z^c z_0 F_1(z) [F_2(z_0)((c-1)\lambda_m - 1 - \omega) - z_0 \lambda_m F_4(z_0)] \\ &\quad + z z_0^c F_2(z) [F_1(z_0)(1 + \omega) + \lambda_m z_0 F_3(z_0)], \end{aligned} \quad (89)$$

with the function definitions

$$\begin{aligned} F_1 &= {}_2F_1(a, b; c; z), \\ F_2 &= {}_2F_1(1 + a - c, 1 + b - c; 2 - c; z), \\ F_3 &= \frac{ab}{c} {}_2F_1(1 + a, 1 + b; 1 + c; z), \\ F_4 &= \frac{(1 + a - c)(1 + b - c)}{2 - c} {}_2F_1(2 + a - c, 2 + b - c; 3 - c; z), \end{aligned} \quad (90)$$

and

$$z = \frac{1}{\bar{M}} \left[1 - \frac{\lambda_m}{\alpha_m(1 - x)} \right], \quad z_0 = \bar{M} z. \quad (91)$$

Here, $\bar{M} \equiv e^{\lambda t}$ and is related to \bar{N} , the average number of wild-type cells, by

$$\bar{N} \equiv e^{\lambda_w t} = \bar{M}^{\lambda_w/\lambda_m} \approx \bar{M}^b. \quad (92)$$

Then, $W_{\bar{N}}(x)$ is given by

$$W_{\bar{N}}(x) = \mathcal{N} \frac{T_1 T_4 - T_2 T_3}{T_4^2} \left(-\frac{T_3}{T_4} \right)^{N-1}. \quad (93)$$

For the calculation of the $T_1 T_4 - T_2 T_3$ factor, it is sufficient to work to zeroth order in μ . To this order $F_1 = F_2 = 1$ and $F_3 = F_4 = 0$. Then, we have

$$\begin{aligned} T_1 T_4 - T_2 T_3 &\approx [bz_0^c z] [z^c z_0 (b\lambda_m - 1) + zz_0^c] - [-bz_0^c z] [z^c z_0 - z_0^c z] \\ &= b^2 \lambda_m z_0^{c+1} z^{c+1}. \end{aligned} \quad (94)$$

For the T_4^2 term in the denominator, things are even simpler since we can use the fact that $z \ll z_0$:

$$T_4^2 \approx z_0^{2c} z^2, \quad (95)$$

so that

$$\frac{T_1 T_4 - T_2 T_3}{T_4^2} \approx b^2 \lambda_m \left(\frac{z}{z_0} \right)^b \approx b \lambda_m / N. \quad (96)$$

For the calculation of T_3/T_4 we have to keep both the first $1/N$ and μ corrections since the ratio is to be raised to the \bar{N} th power. We have, since $a \sim \mathcal{O}(\mu)$

$$\begin{aligned} F_1(x) &\approx 1 + a \sum_{n=1}^{\infty} \frac{\Gamma(b+n)\Gamma(b+1)}{\Gamma(b)\Gamma(b+1+n)n} x^n \\ &= 1 + \frac{a}{b} [1 - b \ln(1-x) - {}_2F_1(1, b; 1+b; x)] \equiv 1 + \mu F_1^1(x), \end{aligned} \quad (97)$$

and

$$\begin{aligned} F_3(x) &\approx a \frac{\Gamma(b+1)}{\Gamma(b)} \sum_{n=1}^{\infty} \frac{n\Gamma(n)\Gamma(b+n)}{n!\Gamma(b+n+1)} x^{n-1} \\ &= a \frac{{}_2F_1(1, b; b+1; x) - 1}{x} \equiv \mu F_3^1(x), \end{aligned} \quad (98)$$

and so

$$T_3 \approx -z_0^c z \left[1 + \mu F_1^1(z_0) - \frac{1}{\bar{N}} \right]. \quad (99)$$

Similarly,

$$T_4 \approx z_0^c z \left[1 + \omega + \mu \lambda_m z_0 F_3^1(z_0) + \mu F_1^1(z_0) + \frac{b\lambda_m - 1}{\bar{N}} \right]. \quad (100)$$

Thus

$$\frac{T_3}{T_4} \approx - \left[1 - \frac{b\lambda_m}{\bar{N}} - \omega - \mu \lambda_m z_0 F_3^1(z_0) \right]. \quad (101)$$

Thus, we have our result

$$W_{\bar{N}}(x) \approx \mathcal{N} \left(\frac{b\lambda_m}{\bar{N}} \right) \exp \left(-b\lambda_m - \bar{N}\omega - \mu \bar{N} \lambda_m z_0 F_3^1(z_0) \right). \quad (102)$$

Since, when $x \rightarrow 1$, $z_0 \rightarrow -\infty$ and $\mu F_3^1(z_0) \rightarrow -a/z_0$, the condition $W_{\bar{N}}(1) = 1$ implies

$$\begin{aligned} W_{\bar{N}}(x) &\approx \exp\left(-\mu\bar{N}\lambda_m z_0(F_3^1(z_0) + a/z_0)\right) \\ &= \exp\left(-\mu N \left(\frac{b_w}{b_m}\right) \left(\frac{b_m - d_m}{b_w - d_w}\right) {}_2F_1(1, b; b+1; z_0)\right). \end{aligned} \quad (103)$$

This recovers the result we obtained above, as expected.

It is amazing how our relatively simple scaling distribution is encoded in the immensely more complicated AK formalism. Being exact, their equation must contain all the details regarding the outer region (when m is the same order as N), must allow for arbitrary values of μ and N , and must reflect the intricate interplay of time and population size. Our approach is much less general, but has the distinct advantage of leading to answers that are simple enough for us to gain some insight into what is going on. In the last part of this section, we will see that we can employ a trick to recover our results from the AK formalism in a much simpler fashion.

5.3 An alternate method

There is an alternate way to extract our desired fixed population distribution from the AK results. We have seen that in the regime of interest, $m \ll N$, the distribution is independent of N_0 . If we take N_0 large however, the wild-type dynamics becomes deterministic and so the fixed time and fixed population ensembles become identical. Thus, all we need to do is to take the AK results for large N_0 , and this gives the desired generating function directly. We now proceed to carry this out for the general non-symmetric case.

Following the notation in AK, we focus on

$$W = 1 + \omega + \lambda_m \Psi(z), \quad (104)$$

with the definition

$$\Psi(z) \equiv \frac{z^g F_3(z) + C(1-g)F_2(z) + CzF_4(z)}{z^{g-1}F_1(z) + CF_2(z)}, \quad (105)$$

where the F 's are the same as given above, with the constants taking the values (valid to leading order in μ):

$$\begin{aligned} \omega &\approx \mu \frac{b_m - d_m}{b_w - d_w} \frac{b_w}{b_m}; & \lambda_m &\approx \frac{b_m - d_m}{b_w}; \\ e &\approx \mu \frac{b_w^2}{b_m(b_w - d_w)}; & f &\approx \frac{b_w - d_w}{b_m - d_m}; & g &\approx d + 1, \end{aligned} \quad (106)$$

and where the constant C is determined by the normalization condition $W(t=0) = y$:

$$C = z_0^g \frac{\kappa F_1(z_0) - F_3(z_0)}{(1-f-\kappa z_0)F_2(z_0) + z_0 F_4(z_0)}, \quad (107)$$

with z_0 the value of z at $t = 0$. Note that in general z is proportional to $e^{\lambda_m t}$. Finally $\kappa \equiv -\frac{\omega}{\lambda_m z_0} \approx -\frac{e}{z_0}$. This seems to be an completely unwieldy mess. But, we get an enormous simplification if we recall that we are interested in small μ and therefore can use the leading order results $F_1 = 1$, $F_2 = 1$, and F_4 is $O(\mu)$ and hence can be dropped in the denominator, since C is also $O(\mu)$. with the definition $F_3^1(z) = F_3^1(z)/\mu$, we can derive the much simpler leading order expression

$$C \approx z_0^g \frac{e/z_0 + \mu F_3^1(z_0)}{f} \equiv \mu C_1. \quad (108)$$

Finally, we use the fact that z is small as it is related $e^{-\lambda_m t}$ which equals N_0/N raised to the positive power f . Thus, $W(x)$ is given by

$$W(x) = (1 + \lambda_m \Psi(z))^{N_0} = \exp \left(-\mu N \left(\frac{b_m - d_m}{b_w} \right) (d/\mu + z_0 F_3^1(z_0)) \right). \quad (109)$$

We need an explicit expression for $F_3^1(x)$. We have

$$\begin{aligned} F_3^1(x) &= \frac{a}{\mu} \frac{\Gamma(b+1)}{\Gamma(b)} \sum_{n=1}^{\infty} \frac{n \Gamma(n) \Gamma(b+n)}{n! \Gamma(b+n+1)} x^{n-1} \\ &= \frac{ab}{\mu} \sum_{n=1}^{\infty} \frac{x^{n-1}}{b+n} \\ &= \frac{a}{\mu} \frac{{}_2F_1(1, b; b+1; x) - 1}{x}. \end{aligned} \quad (110)$$

Thus, we have the final expression

$$\begin{aligned} W(x) &= \exp \left(-\mu N \left(\frac{b_m - d_m}{b_w} \right) \left(\frac{b_w^2}{b_m(b_w - d_w)} \right) {}_2F_1(1, b; b+1; z_0) \right) \\ &= \exp \left(-\mu N \left(\frac{b_w}{b_m} \right) \left(\frac{b_m - d_m}{b_w - d_w} \right) {}_2F_1(1, b; b+1; z_0) \right), \end{aligned} \quad (111)$$

again reproducing what we have previously calculated.

6 Conclusions

We have analyzed here the pure birth as well as the birth-death stochastic Luria-Delbrück model in the limit of small mutation rate μ and large current population, N , where the product μN is unrestricted. The answer in all cases exhibits an intermediate asymptotic power-law decay m^{1+r} for $1 \ll m \ll N$, where r is the ratio of the wild-type net growth rate to the mutant net growth rate. For small μN , the distribution is monotonically decreasing, and approaches the pure power-law already for m 's of order unity. As μN increases the probability of a small number of mutants decreases sharply, and above a critical μN of order 1, the distribution develops a peak. The power-law decay sets in then only a distance past this peak. For large μN ,

the distribution becomes the one-sided Lévy distribution with index r . The distribution for $m \ll N$ is in all cases independent of the initial wide type colony size N_0 . It is also independent of the precise nature of the ensemble, as long as there are sufficient numbers of wild-type individuals in all members of the ensemble with non-vanishingly small weighting. This can be guaranteed by fixing $N = \bar{N}$ at either fixed time or in a time average, or alternatively by having a large number for the number N_0 of initial wild-type individuals. The power-law tail, as well as the strong dependence of the distribution on N_0 for $m \sim N$ has significant implications for tasks such as using measured data to infer mutation rate [10, 7]. This also has implications for the variation expected in patient responses to administered therapy, although it seems that the *extrinsic* variation due to different rates dominates over the *intrinsic* variation due to process stochasticity [12].

Acknowledgements This work was supported by the NSF Center for Theoretical Biological Physics, (grant no. PHY-1308264). HL was also supported by CPRIT Scholar program of the State of Texas, and DK was also supported by the Israeli Science Foundation.

References

1. Angerer, W.P.: An explicit representation of the luria-delbrück distribution. *J. Math. Biol.* **42**(2), 145–174 (2001)
2. Antal, T., Krapivsky, P.: Exact solution of a two-type branching process: models of tumor progression. *J. of Stat. Mech., Theory and Experiment* **2011**(8), P08,018 (2011)
3. Clayton, E., Doupe, D.P., Klein, A.M., Winton, D.J., Simons, B.D., Jones, P.H.: A single type of progenitor cell maintains normal epidermis. *Nature* **446**(7132), 185–189 (2007)
4. Coldman, A.J., Goldie, J.H.: A stochastic model for the origin and treatment of tumors containing drug-resistant cells. *Bull. Math. Biol.* **69**(3), 279–292 (1986)
5. Dewanji, A., Luebeck, E.G., Moolgavkar, S.H.: A generalized luria-delbrück model. *Math. Biosci.* **197**(2), 140–152 (2005)
6. Diaz Jr., L.A., Williams, R.T., Wu, J., Kinde, I., Hecht, J.R., Berlin, J., Allen, B., Bozic, I., Reiter, J.G., Nowak, M.A., Kinzler, K.W., Oliner, K.S., Vogelstein, B.: The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* **486**(7404), 537–540 (2012)
7. Gerrish, P.: A simple formula for obtaining markedly improved mutation rate estimates. *Genetics* **180**(3), 1773–1778 (2008)
8. Iwasa, Y., Nowak, M.A., Michor, F.: Evolution of resistance during clonal expansion. *Genetics* **172**(4), 2557–2566 (2006)
9. Kashdan, E., Pareschi, L.: Mean field mutation dynamics and the continuous luria-delbrück distribution. *Math. Biosci.* **240**(2), 223–230 (2012)
10. Kepler, T.B., Oprea, M.: Improved inference of mutation rates: 1. an integral representation for the luria-delbrück distribution. *Theor. Pop. Biol.* **59**(1), 41–48 (2001)
11. Kessler, D.A., Levine, H.: Large population solution of the stochastic luria-delbrück evolution model. *Proc. Nat. Acad. Sci.* **110**(29), 11,682–11,687 (2013)
12. Kessler, D.A., Levine, H., Austin, R.: Resistance to chemotherapy: Patient variability and cellular heterogeneity. *Cancer Res.* (in press)
13. Komarova, N.: Stochastic modeling of drug resistance in cancer. *J. Theor. Biol.* **239**(3), 351–366 (2006)
14. Lea, D.E., Coulson, C.A.: The distribution of the numbers of mutants in bacterial populations. *J. Genetics* **49**(3), 264–285 (1949)

-
15. Lenski, R.E., Slatkin, M., Ayala, F.J.: Mutation and selection in bacterial populations: alternatives to the hypothesis of directed mutation. *Proc. Nat. Acad. Sci.* **86**(8), 2775–2778 (1989)
 16. Luria, S.E., Delbrück, M.: Mutation of bacteria from virus sensitivity to virus resistance. *Genetics* **28**(6), 491–511 (1943)
 17. Mandelbrot, B.: A population birth-and-mutation process, i: Explicit distributions for the number of mutants in an old culture of bacteria. *J. Appl. Prob.* **11**(3), 437–444 (1974)
 18. Moolgavkar, S.H.: Carcinogenesis modeling: from molecular biology to epidemiology. *Ann. Rev. Publ. Health* **7**(1), 151–169 (1986)
 19. Tomasetti, C., Levy, D.: An elementary approach to modeling drug resistance in cancer. *Math. Biosci. Eng.* **7**(4), 905–918 (2010)
 20. Zheng, Q.: Progress of a half-century in the study of the luria-delbrück distribution. *Math. Biosci.* **162**(1), 1–32 (1999)